(4)

ANNUAL REPORT

FONR OPTICAL COMPUTING ST

N00014-86-K-0591
5-31538

Prepared By

H. J. Caulfield
Center for Applied Optics
University of Alabama in Huntsville
Huntsville, Alabama 35899

SEPTEMBER 13, 1988

ANNUAL REPORT

I. INTRODUCTION

As noted in our original proposal, most of our effort has been in hybrid analog-digital algebra and in massively parallel processing based on hologram arrays. Both programs have exceeded our announced goals as will become clear from what follows. In the process of those primary studies, we developed a number of other worthwhile applications in optics as well.

Future work will aim almost exclusively at the holographic interconnect as this is likely to be of most immediate value in meeting announced SDI/ONR future program needs and we have reached a satisfactory conclusion to the algebra study by laboratory demonstration and theoretical justification of the results previously obtained only by computer simulation and justifying arguments.

This report is comprised of overview and detail parts for each application. The detail is relegated to appendices to make reading of the overview more convenient.

II. OVERVIEW

2.1 HOLOGRAPHIC N^4 INTERCONNECT

Despite the widespread belief that N^4 interconnect is impossible, it has been done for many years. Fourier optics connects each of NXN input pixels to each of NXN output (Fourier transform) pixels. What we have sought so far is a fixed full rank N^4 interconnect matrix, i.e. N^4 fully independent weighted interconnection paths. When N reaches the range of 100 to 1000, this is more parallel interconnections than electronics will ever achieve and, therefore, establishes a unique niche for optics. The probable application is neural networks.

Peter Guilfoyle for Opticomp has noted that same unique niche for optics in digital computing. The fixed interconnect holograms developed during this reporting period do not serve that need. Accordingly, it is a goal for the next period to make active N^4 interconnects.

The highlight of this reporting period was a careful examination of the inherent and current-technological constraints on N^4 interconnection using fixed holograms. We showed (Appendix A) that with currently available components we could make $(256)^4 \approx 4 \times 10^9$ parallel interconnections. With redesigned components, we could make $(1024)^4 \approx 10^{12}$ parallel interconnections.

2.2 OPTICAL ALGEBRA

According to various estimates somewhere between 50% and 75% of all CPU time in the United States is spent in solving some sort of <u>linear algebra</u>. Examples include least squares analysis, antenna beam steering, linear regression, computational fluid dynamics, finite element analysis, or simply N linear equations with N unknowns.

Other <u>nonlinear algebra</u> problems are also important. These include image processing, linear programming, and super resolution.

To the extent that optics can solve such problems in a parallel fashion, it can lead to small, fast processors which would greatly improve the utility of trackers, radar, sonar, etc.

<u>WHAT IS THE CURRENT STATUS?</u>

We want to solve problems like

$$2x_1 + 3x_2 + x_3 = 4$$

$$3x_1 + x_2 + 3x_3 = 2$$

$$3x_1 + 4x_2 + 7x_3 = 1$$

We can represent these generally as

$$\vec{A}\,x = \vec{b}.$$

In this case

$$A = \begin{bmatrix} 2 & 3 & 1 \\ 3 & 1 & 3 \\ 3 & 4 & 7 \end{bmatrix}$$

$$\vec{b} = \begin{bmatrix} 4 \\ 2 \\ 1 \end{bmatrix} \quad \text{and} \quad \vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

The matrix A and the vector $\vec{b}$ are given. We seek the vector $\vec{x}$.

There is a way to assign a single number (a "norm") to vectors and matrices. We normally use the Euclidean norm, e.g.

$$\| \vec{x} \| = \left[x_1^{\,2} + x_2^{\,2} + x_3^{\,2} \right]^{1/2}$$

The word "solve" has two different meanings. We presume there is a "true" answer x_T. We can say we have an ϵ-accurate solution if

$$\| \vec{x} - \vec{x}_T \| < \epsilon.$$

A weaker sense of "solve" is

$$\| \vec{b} - A\,\vec{x} \| < \epsilon.$$

This is weaker in the rough sense that some good solutions in this sense may not be close to $\vec{x}_T$. On the other hand, for many problems, this "low residual" solution is perfectly adequate. The Bimodal Optical Computer (BOC) minimizes the residual.

One speaks of computational complexity in terms of how something scales with some resource. We will speak of spatial and temporal complexity. We will represent an NXN matrix in parallel using N^2 numbers. We say the spatial complexity scales on the order of N^2, written $O(N^2)$. We will show that the temporal complexity is O(1), i.e., independent of N, provided that N is small enough to be represented spatially in our processor.

The most basic concepts are over a century old (due to Lord Kelvin).

(1) We use a fast, low-accuracy processor to obtain a first guess $\vec{x}_0$,

(2) We use a slow, accurate processor to evaluate the residual

$$\vec{r}_0 = \vec{b} - A\,\vec{x}_0,$$

If $\| \vec{r}_0 \| < \epsilon$, stop.

(3) Otherwise, use the low accuracy solver to

solve for $\Delta\,\vec{x}_0 = \vec{r}_0$.

If we could solve that problem accurately, then

$$\vec{x}, = \vec{x}_0 + \Delta\,\vec{x}_0$$

would have zero residual.

Thus

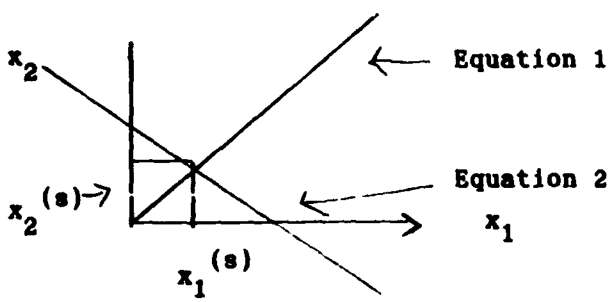
$$\begin{aligned} A\,\vec{x}_1 &= A\,(\vec{x}_0 + \Delta\,\vec{x}_0) \\ &= A\,\vec{x}_0 + A\,\Delta\,\vec{x}_0 \\ &= A\,\vec{x}_0 + \vec{r}_0 \\ &= A\,\vec{x}_0 + \vec{b} - A\,\vec{x}_0 \\ &= \vec{b}. \end{aligned}$$

(4) Use the slow, accurate processor to evaluate

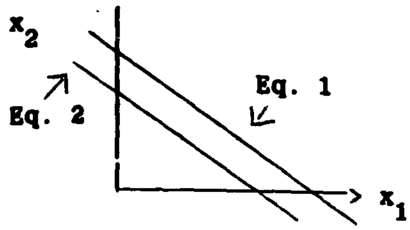
$$\vec{r}_1 = \vec{b} - A\,\vec{x}_1$$

If $\| \vec{r}_1 \| < \epsilon$, stop. Otherwise go to (3).

Some algebra problems resist accurate solution more than others. In high school we solved N=2 problems graphically.



The solution is $x_1^{(s)}$, $x_2^{(s)}$. Problems like this are said to be "well conditioned" and are quite rare in real life. A more common case is



Such problems are said to be "ill conditioned." If the lines are parallel, we say A is "singular." Let us now make this somewhat more rigorous. Let us define a "condition number"

$$\chi(A) = \| A \| \cdot \| A^{-1} \|.$$

Then

$$\epsilon(\| \vec{x} \|) = \chi(A)\, \epsilon(P),$$

where

$\epsilon(\| \vec{x} \|)$ = relative error in the result and

$\epsilon(P)$ = relative accuracy of the processor.

If we have $\epsilon(P) = 0.1$ (very good optics) and $\chi(A) = 10$ (wonderfully benign problem),

$$\epsilon(\| \vec{x} \|) = 1,$$

i.e., 100% errors are likely.

This why we go to 32 bit floating point electronics. No one wants an answer accurate to one part in 2^{32} ($\sim 4 \times 10^9$). We need that to get meaningful answers for large χ. The ultimate ill-conditioning, singularity, corresponds to infinite χ. Such problems are common.

In roughly 1985, Caulfield showed that this iterative process converges (roughly) if

$$\epsilon(P) < \frac{1}{2\chi(A)}.$$

For good optics, ϵ (P) = 0.1. Thus we need

$$\chi (A) < 5$$

to guarantee solution. This is silly. No real problems are so benign.

In 1987 we showed that replacing A by $A' = A + E$ where E is an error matrix and

$$\| E \| / \| A \| \ll 1,$$

leads to convergence for all problems independently of χ. For large χ, the x which minimizes $\| \vec{r} \|$ may be less close to $\vec{x}_T$ than would be the case for small χ. Nevertheless, we can drive $\| \vec{r} \|$ to zero in very few iterations even for singular matrices. Call this Breakthrough 1.

To do the fast, low-accuracy solution 0 (1) in time; we use another trick. We employ a parallel $A \vec{x} = \vec{y}$ device.

$$\begin{array}{ccccc} x_1 & \longrightarrow & \boxed{} & \longrightarrow & y_1 \\ x_m & \longrightarrow & \boxed{A} & \longrightarrow & y_m \end{array}$$

These are easy in optics. Wai Cheng and Caulfield showed that if we correct x_k with a signal proportional to $b_k - y_k$, for all k, then this system would "relax from any starting $\vec{x}$ to one satisfying $A \vec{x} = \vec{b}$ (in the low $\| \vec{r} \|$ sense) under the circumstance that A is "positive definite." To explain this, we need one more diversion.

A vector e such that

$$A \vec{e} = \lambda \vec{e},$$

where λ is a scalar, is said to be an "eigenvector" of A. We usually normalize $\vec{e}$, i.e. set

$$\| \vec{e} \| = 1.$$

In that case, λ is the corresponding "eigenvalue." Let us arrange the eigenvalues of A such that

$$\lambda_1 < \lambda_2 < \cdots < \lambda_r$$

(r connotes "rank," a concept we choose not to define here). Interestingly,

$$\chi(A) = |\lambda_r|/|\lambda_1|.$$

The interesting thing for our purposes is that the relaxation processor converges at a rate (roughly) of

$$e^{-\lambda_1 t}.$$

Obviously if $\lambda_1 > 0$, it does not converge. A matrix for which $\lambda_1 > 0$, is said to be positive definite. A matrix $B = \begin{bmatrix} 12 \\ 34 \end{bmatrix}$ can undergo a row-for-column switch to form a transpose

$$B^T = \begin{bmatrix} 13 \\ 24 \end{bmatrix}.$$

Since the matrix elements may be complex, we can complex conjugate a matrix A to get A^*. We call

$$(A^*)^T = (A^T)^* = A^H,$$

the Hermitian of A. For any matrix both AA^H and $A^H A$ are nonnegative definite ($\lambda_1 > 0$). We noted that $A^H A + E$ and $AA^H + E$ are positive definite if $\bar{E} > 0$.

Note, though,

$$A \vec{x} = \vec{b}$$

$$A^H A \vec{x} = A^H \vec{b}.$$

Write

$$B = A^H A$$

and

$$\vec{c} = A^H \vec{b}.$$

Then

$$B \vec{x} = \vec{c}$$

and B is nonnegative definite (likewise for AA^H). Applying our method to this makes all methods converge even though

$$\chi(A^H A) = \chi(AA^H) = \chi^2(A),$$

a normally-disastrous event. These realizations are Breakthrough 2.

Many other things done in BOCs are pretty, but those two are essence. Of the two, Breakthrough 1 is essential. Breakthrough 2 allows 0 (1) solutions.

SUMMARY

CONVENTIONAL ALGEBRA ON DIGITAL COMPUTERS	BIMODAL OPTICAL COMPUTERS
• SEEKS $\Vert \vec{x} - \vec{x}_T \Vert < \epsilon$	• SEEKS $\Vert \vec{b} - A\vec{x} \Vert < \epsilon$
• REQUIRES ROUGHLY $O(N^3)$ TEMPORAL COMPLEXITY	• 0 (1) TEMPORAL COMPLEXITY
• ALGORITHM MATCHED TO PROBLEM	• CONSTANT ALGORITHM SUFFICES
• $\epsilon(\Vert \vec{x} \Vert) \propto \chi(A)$	• $\Vert \vec{b} - A\vec{x} \Vert \rightarrow 0$ INDEPENDENTLY OF $x(A)$
• $E(\Vert \vec{x} \Vert) \propto E(P)$	• $\Vert \vec{b} - A\vec{x} \Vert < \epsilon$ INDEPENDENTLY OF $E(P)$

The highlights of this period include a laboratory demonstration of an 0 (1) time solver of even singular matrix equations and the first vigorous mathematical proof of how this works. Appendix B gives those details.

2.3 PATTERN RECOGNITION

In an early part of this contract we showed that rotation invariant pattern recognition masks much simpler to make than those of Arsenault and much simpler to use than those of Sweeny could be made. These were simply annular rings of fixed amplitude and phase. In Appendix C we show experimental work which shows that these simple filters actually outperform their more complex competitors.

2.4 RESIDUE ARITHMETIC

In an early phase of this work we showed that 2D and 3D optical interconnect matrices which we called optical Fredkin gates offer some real advantages over other interconnection arrays. In Appendix D we show that optical Fredkin gates make possible much higher speed residue arithmetic calculation chains than any method so far proposed.

III. CONCLUSION OF THIS PERIOD

Our original goals on optical algebra have all been met or exceeded. The basic work, including theory and laboratory demonstration, has been completed.

The concepts originated in the prior reporting period for improvements in optical pattern recognition masks and optical Fredkin gate arrays have been carried to the points where they show demonstrable advantages over prior methods.

The fixed holographic N^4 interconnection system establishes that optics can make 4×10^9 weighted independent interconnections now and 10^{12} eventually. These are tasks of great interest in neural networks but well beyond current or projected electronics capability.

Future work will concentrate on applications of holographic N^4 interconnections to neural networks and to Peter Guilfoyle's digital computer.

APPENDIX A

HOLOGRAPHIC INTERCONNECTIONS

1. "Parallel N^4 Weighted Optical Interconnections," H. J. Caulfield, Applied Optics, Vol. 26, No. 19 (October 1987), pg. 4039.

2. "The Holographic Basis for Intelligent Machines," H. J. Caulfield in John Robillard, Editor, Practical Holography, to be published by Oxford University Press (1989).

3. "Massive Holographic Interconnection Networks and Their Limitations," H. J. Caulfield, R. Barry Johnson and Joseph Shamir, submitted to Applied Optics.

APPENDIX B

OPTICAL ALGEBRA

1. "On An Iterative Method for Consistent Linear Systems," Peter M. Gibson and H. J. Caulfield, submitted to Linear and Multi-Linear Algebra.

2. "Superconvergence of Hybrid Optoelectronic Processors," Mustafa Abushagur, H. J. Caulfield, Peter M. Gibson, and Mohammad Habli, Applied Optics, Vol. 26, No. 23, pg. 4906-4906 (December 1987).

3. "Adaptive Array Radar Data Processing Using the Bimodal Optical Computer," Mustafa Abushagur, Microwave and Optical Technology Letters, Vol. 1, No. 7 pg. 236-240 (September 1988).

4. "Jam Resistance of the Bimodal Optical Computer," Mustafa Abushagur and H. John Caulfield, SPIE Proceedings, Vol. 886, pg. 171-178 (1988).

5. "Hybrid Optoelectronic Nonlinear Algebra Processor," Mustafa Abushagur and H. John Caulfield, SPIE Proceedings, Vol. 936, pg. 309-314 (1988).

6. "Solving System of Linear Equations Using The Bimodal Optical Computer (Experimental Results), Mustafa Abushagur, H. John Caulfield, and M. A. Habli, SPIE Proceedings, Vol. 936, pg. 315-320 (1988).

7. "Solving Ill-Posed Algebra Problems Using The Bimodal Optical Computer," Mustafa Abushagur and H. John Caulfield, SPIE Proceedings, Vol. 939, pg. 29-33 (1988).

8. "Real-Time Optical Processing Of Antenna Array Signals," Joseph Shamir, H. John Caulfield and Brian M. Hendrickson, Microwave and Optical Technology Letters, Vol. 1, No. 3, pg. 100-102, (May 1988).

APPENDIX C

PATTERN RECOGNITION

1. "Pattern Recognition Using Reduced Information Content Filters," H. J. Caulfield, Joseph Shamir, and Joseph Rosen, Applied Optics, Vol. 26, No. 12, pg. 2311-2314 (June 1987).

2. "High-efficiency Rapidly Programmable Optical Interconnections," H. J. Caulfield and Joseph Shamir, Applied Optics, Vol. 26, pg. 1032-1037 (1987).

3. "Distortion Invariant Pattern Recognition with Phase Filters," J. Shamir and H. J. Caulfield, Applied Optics, Vol. 26, pg. 2315-2319 (1987).

4. "Circular Harmonic Phase Filters for Efficient Rotation Invariant Pattern Recognition," J. Rosen and J. Shamir, Applied Optics, Vol. 27, pg. 2895-2899 (1988).

APPENDIX D

RESIDUE ARITHMETIC

1. "Three-Dimentional Optical Interconnection Gate Array," Joseph Shamir, Applied Optics, Vol. 26, pg. 3455-3457 (1987).

2. "Residue Arithmetic Processing Utilizing Optical Fredkin Gate," Applied Optics, Vol. 26, pg. 3941-3946 (1987).

3. "Optical Interconnection Network Using Polarization-based Ferroelectric Liquid Crystal Gates," K. M. Johnson, M. R. Surette and J. Shamir, Applied Optics, Vol. 27, pg. 1727-1733 (1988).